

# Lexical Items in Complex Predications: Selection as Underassociation

Gillian Ramchand  
*University of Tromsø, CASTL*

## Abstract

This paper examines the problem of selectional ‘matching’ effects in Bengali V-V complex predicates, and English denominal verbs within the context of a compositional syntax/semantics for verbal meaning and a theory of lexical insertion under non-terminals. It argues that within the particular version of this kind of lexical insertion, as proposed by Ramchand 2008b, selection can be captured by the underassociation of category features constrained by AGREE. In this way, I argue that we can achieve many of the effects of selection without any distinct lexical subcategorization frame, or sub-type of feature-checking, once we have a suitably articulated theory of lexical insertion.

## 1. Introduction

In the past, discussions of argument structure were rooted in a view of grammar that had the Lexicon as an important locus for thematic and selectional information. Selectional information came in two main types: c-selection and s-selection (categorical selection and semantic selection respectively). The former type was classically encoded in subcategorization frames. The latter is essentially a grabbag of semantic features, thematic role information, or even real world felicity conditions which were argued to be outside of grammar.

- (1) SELECTION BASED ON A LEXICAL MODULE (WITH A ‘PROJECTION PRINCIPLE’)
  - a. *c-selection*:  
‘Matching’ via subcategorization frames
  - b. *s-selection*:  
‘Matching’ via semantic features; thematic role specifications; real world knowledge

---

\* Thanks to Peter Svenonius and Rajesh Bhatt for many discussions of the ideas presented in this paper. Portions of this paper have been presented at various conferences over the past few years. I thank the audiences of the Basque Country conference on Argument Structure in May 2007, the NORMS conference on argument structure in Lund, February 2008 and the ‘Words don’t come easy’ conference in Verona in November 2008. More generally, the ideas in this paper have emerged in the context of the fruitful and stimulating discussions at CASTL and the University of Tromsø. Since I can no longer disentangle all their different contributions, I simply assert that they must be all of the good things, and none of the bad things.

In a world view where the Lexicon does not contain independent rules and primitives and where argument structure is increasingly being understood constructionally, the problem of selection reemerges, with no natural place to locate it. The general strategy in the minimalist literature is to reduce selection to either category selection, or checking of some feature with semantic import (whatever we allow ourselves to justify as being syntactically active). A subset of ‘category selection’ can also be seen to follow from the rigid ordering of the functional sequence (Fseq), or whatever *that* turns out to follow from. Non-syntactically relevant semantic selection should be relegated to real world knowledge and felicity conditions in such a model. I also will assume that some things that have been called ‘selection’ should be so relegated.

(2) SELECTION IMPLEMENTED IN THE SYNTAX

a. *c-selection*:

- (i) Ordering imposed by extended projection/functional sequence
- (ii) Checking of ‘selectional’ category feature at D-structure (or under ‘first merge’). (See Adger 2003)

b. *s-selection*:

- (i) Checking of syntactic features corresponding to certain semantic requirements such as: thematic role feature, animacy, volitionality, telicity etc.

There are of course many different ways to distribute the work done by selection, depending on the one’s framework. One important architecture that deserves a few words here is that of Distributed Morphology (DM). Within DM, the selectional problems I highlight here do not take centre stage in the most recent research, where a lot of emphasis has been placed on a different kind of selection—selection of vocabulary items under competition for insertion under terminal nodes. This type of selection is constrained by the subset principle but can only deal with co-occurrence of features that have been unified under terminal nodes by syntactic operations of morphological ‘fusion’. However, within this architecture it is also possible to encode information about insertion ‘frames’ directly associated with vocabulary items (see Harley and Noyer 1999 for general discussion) which essentially takes over much of the work of lexical subcategorization frames. It is important to realise that the use of insertion frames associated with vocabulary items is essentially a return to the lexicalist use of subcategorization frames, but in the context of late insertion. It is commonly assumed that the memorization of idiomatized chunks of structure (as in Marantz 2001) is also independently required which takes care of another part of the job traditionally captured by selection. On the other hand, certain extreme constructivist views which reject the use of ‘insertion frames’ are forced to deny the fact of selection at all (cf. Borer 2005), but this, I

believe is not tenable. The difficulty and centrality of the selection problem has been, to my mind, underappreciated and underestimated in these models, especially in the more constructivist literature (see also Emonds 2000 for criticism along these lines).

In this paper I will look at some important cases of selection in the domain of complex predication, but attempt to do without (2a<sub>ii</sub>) and (2b<sub>i</sub>), a necessity imposed by the following assumptions.

#### Assumptions:

- The inventory of functional heads within each major category is far larger, more finely articulated and more sensitive to semantic compositionality than in earlier stages of the theory (cf. also Cinque 1999).
- Thematic role labels in the lexicon should be abandoned, as in the common constructionalist view of argument structure, and there are no equivalent syntactic features bearing these labels to be checked, and no listed ‘insertion frames’.
- Selection can be achieved via category labels on lexical items *plus* a constrained theory of lexical insertion (which we need anyway).

I wish to show in this paper that these assumptions lead to a more elegant and satisfying theory, but we will have to give up certain cherished habits concerning lexical insertion.

Essentially, the idea will be that a proper description of the categories lexicalizable by a particular item will derive its proper contexts of insertion. Consider an analogy to the description of the chemical elements in physical chemistry (here, our ‘lexical items’). In such a situation, listing a description of the element’s properties together with an independent listing of the different kinds of element it can combine with would seem like missing a generalization. Instead, our chemical descriptions of elements in terms of protons, neutrons and electrons are successful precisely because they can predict what kinds of chemical compounds they can participate in. I will pursue a similar strategy here with lexical items: the clutch of category features that they are listed with can have predictive power with respect to their combination with other items in complex predication structures.

In exploring the issue of selection, I will use two distinct test cases where matching effects are pervasive: (i) V-V complex predicates in Bengali, and (ii) denominal verbs in English. I hope to show that a proper understanding of the nature of lexical insertion, and in particular the phenomenon I will call ‘underassociation’, provides us with a novel and theoretically satisfying way to account for some of the effects of selection, without undermining the minimalist view of a syntactic lexicon.

## 2. Background

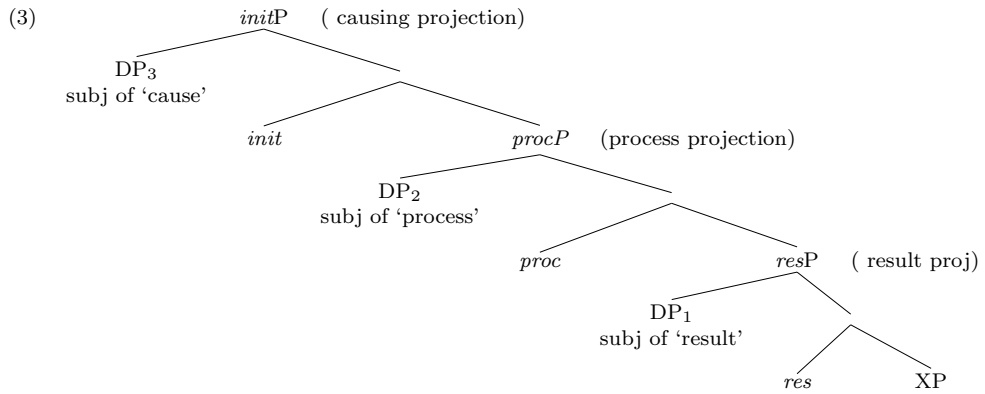
There are two aspects of my background assumptions that need to be spelled out in order to proceed. First of all, I lay out the specifics of the functional decomposition that I assume when it comes to verbs, essentially describing the system in Ramchand (2008b). The details of the decomposition may vary as the theory advances, however, and not affect the general point that I will try to make concerning selection. It is presented here for concreteness.

The second aspect of my background assumptions concerns the idea of lexical insertion. In the second subsection, I lay out a theory of vocabulary item insertion that does not assume that VIs are inserted under terminal nodes. I present a principle called the Superset Principle, argued for by Caha (2007), and based on ongoing work of Michal Starke, which regulates the choice of competitors in this model. This, together with the Elsewhere Condition, and Exhaustive Lexicalization will be the system that replaces the old ‘insert under terminals’ model.

In the main body of the paper, I will show the empirical effects of this system, arguing that the superset principle should be constrained by AGREE in order to account for matching effects in Bengali V-V complex predicates. Finally I turn to the class of denominal verbs and propose an analysis of them in the new system that avoids some of the paradoxes of conflation (Hale and Keyser 2000).

### 2.1. The Functional Sequence

I assume the following as the subpart of the functional sequence that corresponds to the notion of ‘dynamic verb’ in earlier phrase structure models, after Ramchand (2008b). Each projection corresponds to a subevent with its own predication subject position, and linked by the generalized ‘leads-to’ or ‘cause’ relation.



- *initP* introduces the causation event and licenses the external argument (‘subject’ of cause = INITIATOR)
- *procP* specifies the nature of the change or process and licenses the entity undergoing change or process (‘subject’ of process = UNDERGOER)
- *resP* gives the ‘telos’ or ‘result state’ of the event and licenses the entity that comes to hold the result state (‘subject’ of result = RESULTEE) .

There is one basic primitive rule of event composition in this system, the ‘leads to’ relation. represented here as ‘ $\rightarrow$ ’:

- (4) Event Composition Rule:  
 $e = e_1 \rightarrow e_2$  :  $e$  consists of two subevents,  $e_1$ ,  $e_2$  such that  $e_1$  causally implicates  $e_2$   
 (cf. Hale and Keyser 1993)

There are two general primitive predicates over events corresponding to the basic subevent types as follows:

- (5) a. State( $e$ ) :  $e$  is a state  
 b. Process( $e$ ):  $e$  is an eventuality that contains internal change
- (6) IF,  $\exists e_1, e_2$ [State( $e_1$ ) & Process( $e_2$ ) &  $e_1 \rightarrow e_2$ ] , then by definition Initiation( $e_1$ )
- (7) IF  $\exists e_1, e_2$ [State( $e_1$ ) & Process( $e_2$ ) &  $e_2 \rightarrow e_1$ ] then by definition Result( $e_1$ )

Thematic roles are an epiphenomenon of the fact that DPs in specifier positions of these subevental projections acquire entailments based on being the ‘subject of predication’ of those events. We can give labels to the cluster of entailments that are delivered by the interpretational system by virtue of being in the different specifier positions, as follows.

- (8) a. Subject ( $x$ ,  $e$ ) and Initiation( $e$ ) entails that  $x$  is the INITIATOR of  $e$ .  
 b. Subject ( $x$ ,  $e$ ) and Process( $e$ ) entails that  $x$  is the UNDERGOER of the process.  
 c. Subject ( $x$ ,  $e$ ) and Result( $e$ ) entails that  $x$  is the RESULTEE.

In practice these labels do not correspond neatly to thematic roles in the traditional sense, because I assume that DPs can undergo MOVE and accumulate entailments from more than one such specifier position.

We have seen that certain arguments are related in a one-to-one fashion to the projections corresponding to each subevent— they are the ‘subjects’

or ‘specifiers’ of those projections. However, in addition to the specifier positions, DPs (and indeed PPs and APs) may occupy the complement positions of the verbal heads giving rise to what I call ‘rhemes of process’ and ‘rhemes of result’ respectively. Rhematic material, by definition, will never occur in the specifier position of an eventive head; it will always occur in complement position to an eventive head. RHEMES, and as an important subcase PATHS, do not describe elements that are individuated and predicated over within an event topology, but those that actually co-construct the specific predication property (static or dynamic) that the ‘subject’ is asserted to have. In an extension of the terms of Talmy (1978), for example, the specifiers are ‘Figures’ and complements are ‘Grounds’ in an asymmetrical predication relation (see Ramchand 2008b for further discussion of RHEMES in terms of homomorphism, and for examples of rhematic arguments). RHEMES will be important to the discussion of selection, in section 4, since these elements will provide heads along the complement line, or c-selectional line of the verb.

## 2.2. The Superset Principle

Once we have admitted a more articulated functional sequence, a tension arises with respect to lexical insertion, especially under a system where the syntactic features that are relevant for insertion each correspond to a distinct head in the structure. In Ramchand (2008b), I argued that verbal lexical items come with a *set* of category features, and thus need to ‘multi-attach’ in the structures described above. This idea is consistent with the proposals recently formalized by Starke and Caha, rethinking the conditions of lexical insertion and extending it to nonterminal nodes. I would argue that in a framework such as the verbal decomposition offered here, this rethinking is natural and necessary. An alternative would be to use syntactic head movement, or morphological merger to pre-bundle the necessary features under a particular terminal node, as in Distributed Morphology. Since these strategies simply mimic the effect of insertion in larger chunks of structure, I will follow Caha (2007) in advocating the more direct approach, while reformulating the notion of competitors. The proposal is that the choice of competitors is regulated by a ‘Superset’ principle, instead of the commonly assumed ‘Subset’ principle of Distributed Morphology. It is important to realise that the Superset principle can also be combined with the general ‘Elsewhere condition’ to give the generalized superset principle in (9). If the assumption of insertion under terminals is abandoned, then this principle gives equivalent results to the generalized Subset principle in many cases (see Caha 2007 for discussion).

The Superset Principle is given below, and compared with a more standard notion of the Subset principle, both combined with an Elsewhere condition, as articulated in Caha (2007).

(9) **The Superset Principle**

The phonological exponent of a Vocabulary item is inserted into a node if the item matches all or a superset of the grammatical features specified in the node. Insertion does not take place if the Vocabulary item does not contain all features present in the node. Where several Vocabulary items meet the conditions for insertion, the item containing fewer features unspecified in the node must be chosen.

EQUALS A1 PLUS B

*A1 Minimized Superset Principle*

A Vocabulary item applies iff it specifies a superset of the features of a node

*B Elsewhere Condition*

Let  $R_1$  and  $R_2$  be competing rules that have  $D_1$  and  $D_2$  as their respective domains of application. If  $D_1$  is a proper subset of  $D_2$ ,  $R_1$  blocks the application of  $R_2$  in  $D_1$ . (taken from Neeleman and Szendrői 2006).

(from Caha 2007)

(10) **The Subset Principle**

The phonological exponent of a Vocabulary item is inserted into a morpheme in the terminal string if the item matches all or a subset of the grammatical features specified in the terminal morpheme. Insertion does not take place if the Vocabulary item contains features not present in the morpheme. Where several Vocabulary items meet the conditions for insertion, the item matching the greatest number of features specified in the terminal morpheme must be chosen. (from Halle 1997)

EQUALS A2 PLUS B

*A2 Minimized Subset Principle*

A Vocabulary item applies iff it specifies a subset of the features of a terminal morpheme

*B Elsewhere Condition*

Let  $R_1$  and  $R_2$  be competing rules that have  $D_1$  and  $D_2$  as their respective domains of application. If  $D_1$  is a proper subset of  $D_2$ ,  $R_1$  blocks the application of  $R_2$  in  $D_1$ . (taken from Neeleman and Szendrői 2006).

(from Caha 2007)

However, abandoning terminal node insertion as a constraint has independent payoffs in cases where the two principles make different predic-

tions. Caha (2007) shows that certain cases of paradigmatic alternations in Czech provide such test cases, and that the superset principle is empirically superior. He also shows that in general, the subset principle version of morphology can get equivalent results to the superset principle only at the expense of introducing ad hoc principles of Fusion, Fission, and Readjustment in the morphological component (Halle and Marantz 1993). These are now the mainstay of DM and become increasingly necessary as trees become ‘bigger’, since they are invoked purely to mimic the effects of insertion under non-terminal nodes, as Caha convincingly argues.

Moreover, the superset and the subset principles make different assumptions about the architecture of the grammar and in particular, of the relation between the syntax and the lexicon. As Caha (2007) puts it.

“The (Minimized) Subset Principle allows the spell-out procedure to ignore features of syntax, but not those of the lexicon. In other words, every feature specified in the lexical entry must have a matching feature in the syntactic structure, but not every feature of syntax must be “spelled out” (i.e. targeted by a vocabulary item that makes reference to it). The (Minimized) Superset Principle allows the interface to ignore features of the lexical entries (as not every feature specified in the entry for a given exponent has to have a matching feature in the syntax) but, on the other hand, all syntactic features must be spelled out.”

I believe that the latter position is the correct one, and I will express it in (11) as follows.

- (11) **Exhaustive Lexicalization**<sup>1</sup>  
Every node in the syntactic representation must be identified by lexical content.

### 3. Competing Lexical Items

The significant fact about the *Exhaustive Lexicalization* principle is what it neglects to say concerning the overspecification of vocabulary items. Basically, it does not forbid vocabulary items from inserting even though not all of their features are matched by the syntax. This is what I called *Underassociation* in Ramchand (2008b). I show some cases in which I believe this is precisely what happens. However, superset-ism overgenerates if it is not constrained by other principles of grammar. Let’s look at the cases that do occur, relatively uncontroversially.

---

<sup>1</sup>The name and formulation of this principle emerged from collaborative conversations with Antonio Fábregas. See Fábregas (2007) for extensive discussion of its effects in the domain of Spanish directional complements.



### 3.1. ‘Break Off’

Under a standard analysis of the Verb-particle construction, a particle can be combined with an atelic (activity) verb, to create a bounded/telic eventuality.

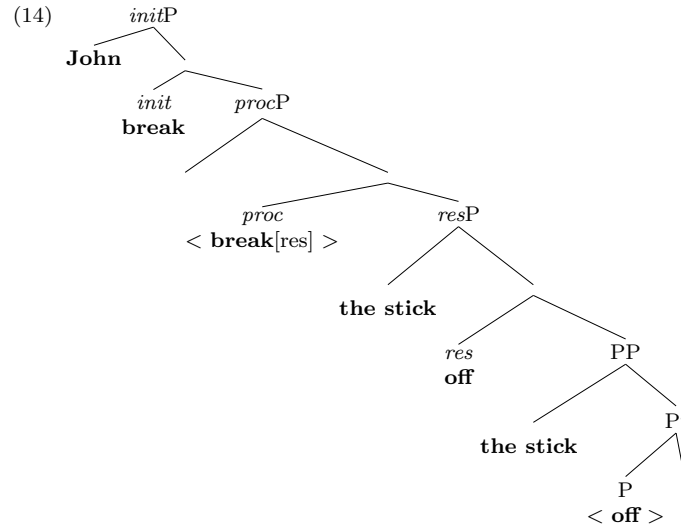
- (12) (a) John pushed the cart for an hour/\*in an hour.  
 (b) John pushed the cart out \*for an hour/in an hour.

Standard analyses, implemented in different ways, would assume that the particle is providing the equivalent of the *[res]* feature here—creating telicity and introducing a small clause predication of result. The activity of ‘pushing’ combined with the result of the cart being ‘out’ create a complex accomplishment predication.

On the other hand, the English verb *break* is telic and resultative in both its transitive and intransitive versions. Let us assume that this means that it has some syntactic feature *[res]* in the system described above. The curious fact is that the *[res]* particle is also grammatical with *break* (*[res]*) verbs.

- (13) (a) John broke the stick in an hour/\*for an hour.  
 (b) John broke the stick off in an hour/\*for an hour.

Both *break* and *off* contribute a resultative feature here, by hypothesis. Exhaustive Lexicalization is satisfied in the tree structure below (essentially following the analysis of Ramchand and Svenonius 2002), but the *[res]* feature of the *break* verb is underassociated (shown here in square brackets).



Note that the result achieved is both a ‘breaking’ and a ‘being off’, so that the lexical encyclopedic content of both items is somehow satisfied. Now let’s look at another example.

### 3.2. Evidence From Bengali

The matching problem I want to discuss comes up with V-V complex predicates in Bengali, where different light verbs are used. Light verb constructions of this type are monoclausal from the point of view of agreement, control and anaphora, and in addition exhibit integrity with respect to scrambling and adverbial modification (see Butt 1995 for a detailed examination of the equivalent construction in Hindi/Urdu). They always give rise to a telic interpretation. The light verb is marked for tense and agreement, and the main verb is explicitly marked as a perfective/conjunctive participle.

There are a number of different light verbs that are used in this construction in Bengali (between 8 and about 16 depending on how one counts), but I will only consider the most common and productive here.<sup>2</sup>

The three light verbs that I will focus on for this discussion are: *jaoya*-‘go’, *oṭha*-‘rise’, *phæla*-‘drop/throw’. My data also includes the extremely common *pOra*-‘fall’ and *tola*-‘lift’, but for the purposes of the ‘matching’ that I found, ‘fall’ patterned like ‘go’ and ‘lift’ patterned like ‘drop’. In the interest of space and clarity, I present only the data from ‘go’, ‘rise’ and ‘drop’ here. All of these light verbs have both a ‘light’ verb use and a ‘heavy’ verb use, conforming to *Butt’s Generalization*.

- (15) *Butt’s Generalization* (Butt 2003, Butt and Lahiri 2005):

Unlike auxiliaries which may become grammaticalized over time to have a purely functional use, light verbs always have a corresponding full or ‘heavy’ version in all the languages in which they are found.

I take this fact to be significant, and will pursue an analysis whereby the categorial feature specification of the ‘light’ verb is no different from the categorial specification that we would have to assume for the ‘heavy’ version. I give some examples below showing the full verb use side by side with light verb use (from Paul 2003).

*jaoya* -‘go’

- (16) (a) *ritu bari gælo*  
           Ritu house go.PAST3  
           ‘Ritu went home.’  
       (b) *gelaš-ṭa bhem-e gælo*  
           glass-CL break-PERFPART go-PAST3  
           ‘The glass broke.’

---

<sup>2</sup>My thanks to Tista Bagchi for her patience and good nature during extensive informant sessions, and for her careful descriptions of contexts of use.

*phæla*-‘drop/throw’

- (17) (a)ritu hat theke boi-ṭa phello  
 Ritu hand from book-CL drop.PAST3  
 ‘Ritu dropped the book from her hand.’  
 (b)ritu kaj-ṭa kor-e phello  
 Ritu work-CL do-PERFPART drop.PAST3  
 ‘Ritu finished her work.’

*Oṭha*-‘rise/climb’

- (18) (a)ritu gačhe uṭhlo  
 Ritu tree-LOC climb.PAST3  
 ‘Ritu climbed the tree.’  
 (b)baloon-ṭa uṭhlo  
 balloon-CLASS rise.PAST3  
 ‘The balloon rose.’  
 (c)ritu hOṭhat heš-e uṭhlo  
 Ritu suddenly laugh-PERFPART rise.PAST3  
 ‘Ritu burst out laughing.’

As can be seen from the data above, *jaoya*-‘go’ and *oṭha*-‘rise’ are intransitive, while *phæla*-‘drop’ is transitive. Moreover, while *jaoya*-‘go’ (and *pOṛa*-‘fall’) pattern like unaccusatives, *oṭha*-‘rise’ seems to be ambiguous between an unaccusative and unergative interpretation.

In classifying these verbs, I will assume the category labels in (19) for transitives, unergatives and unaccusatives. Note that unergatives in the system of Ramchand (2008b) are distinguished by the fact that the specifier of process and the specifier of initiation are identical. Put another way, the initiation head is a ‘raising’ head, and must be filled by internal merge, not external merge. I notate this with a ‘i’ subscript on the relevant head, signifying ‘internal’ Merge.

TRANS: init, proc, (res)  
 UNERG: init<sub>i</sub>, proc, (res)  
 UNACC: proc, (res)

Considering the three light verbs under discussion here, I make the following assumptions about their make-up as full verbs:

*jaoya*-‘go’ is *unaccusative*  
*oṭha*-‘rise/climb’ is ambiguous between an *unergative* and an *unaccusative* reading  
*phæla*-‘drop’ is *transitive*.

I turn now to the selectional facts. While the light verb construction is

ubiquitous in Bengali, and the three light verbs chosen here are extremely productive, there are nevertheless strict selectional restrictions that operate between the particular light verb and the main verb that it combines with. In gathering data on this point, I used a set of about 30 colloquial main verbs drawn from different semantic verb classes and tested them with all of the eight most common light verbs. The data presented here is a small subset of the results, but which conveys the pattern for the three light verbs under consideration here. See Ramchand (2008a) for a more extended write-up of the data, and list of the main verbs tested.

Looking first at the light verb *jaoya*-‘go’, I found that it was exceptionlessly good with intransitive unaccusative verbs as in (19), exceptionlessly bad with unergatives like ‘speak’ and ‘dance’ (20), and also with transitives (21).

- (19) *dorja-ṭa khul-e gælo* *unaccusative*  
 door-CLASS. open-E go.PAST3  
 ‘The door opened.’
- (20) \**O kothaṭa bole gælo* *unergative*  
 he word-CLASS. speak-E go.PAST3  
 ‘He spoke’ (intended)
- (21) \**jon baṛiṭa ban-ye gælo* *transitive*  
 John house-CLASS build-E go.PAST3  
 ‘John built the house’ (intended)

The light verb *pOrā*-‘fall’ behaved the same way, but was only good with a (large) subset of the unaccusative verbs and seemed to require in addition, a sense of suddenness and/or unpredictability that was absent with ‘go’. I put aside the encyclopedic connotations of the light verbs here, and concentrate on the most productive exemplars of each group.<sup>3</sup>

The light verb *phæla*-‘throw/drop’ (and its less productive counterpart *ṭola* -‘lift’ was surprising in that although it is obligatorily transitive, it freely occurred with unergatives (23) in addition to transitives (24), and only systematically rejected the unaccusative verbs that were good with ‘go’ and ‘fall’ (22). The transitive light verb *ṭola*-‘lift’ matched this pattern but with much reduced productivity. The informant seemed to want the result state of the complex predicate formed with ‘lift’ to be a desired final state in some way.

- (22) \**dorja-ṭa khul-e phello* *unaccusative*  
 door-CLASS. open-E throw/drop.PAST3  
 ‘The door opened.’ (intended)

<sup>3</sup>There is also a more ‘aspectual’ use of *jaoya*-‘go’ as a light verb which has quite distinct semantics along the lines of ‘went on V-ing’, and on this meaning it was able to combine with verbs of all types and valency. I put this reading aside here.

- (23) ram bol-e phello *unergative*  
 Ram speak-E throw/drop.PAST3  
 ‘Ram blurted something out.’
- (24) O bari-ṭa ban-iye phello *transitive*  
 he/she house-CLASS. build-E throw/drop.PAST3  
 ‘He accomplished/completed the building of the house.’

Finally, the light verb *oṭha*, was the most forgiving of the four we have seen so far. It seemed to combine productively with main verbs in both the unaccusative and unergative classes (25) and (26).

- (25) boroph jom-e uṭhlo *unaccusative*  
 snow/ice freeze/accumulate-E rise-PAST.3RD  
 ‘The ice/snow accumulated.’
- (26) O bol-e uṭhlo *unergative*  
 he/she speak-E rise.PAST3  
 ‘He/she spoke up.’

While the constraints on it are not clear, some transitives were bad with *oṭha* ‘rise’ (27), but I found a few that were good (28), something that was robustly out of the question for the other two intransitive light verbs above. In the case of (28), my informant could not distinguish between its meaning and the meaning of the same sentence using *phæla* ‘drop’ as a light verb instead.

- (27) \*O bari-ṭa ban-iye uṭhlo *transitive*  
 he/she house-CLASS. build-E rise.PAST3  
 ‘He accomplished/completed the building of the house.’ (intended)
- (28) O bagh-ṭa-ke lathi mer-e uṭhlo *transitive*  
 he/she tiger-CLASS-ACC kick hit-E rise.PAST3  
 ‘He unexpectedly kicked the tiger.’

To summarise the results of this section, a light verb in Bengali imposes a requirement on the argument structure of the main verb it combines with (independent of any other additional semantic constraints it may impose).

#### Generalization:

- *jaoya* ‘go’ (and *poṛa* ‘fall’) can only combine with unaccusative main verbs.

- *phæla*-‘drop’ (and *tola*-‘lift’) can combine only with transitive and unergative main verbs.
- *oṭha*-‘rise/climb’ can combine with either unaccusative or unergatives, and even one or two transitives.

If we look at the category specifications assumed in Ramchand (2008b) for independent reasons, we find that the light verbs and the main verbs seem to need to ‘match’ in Bengali, assuming that the light verb is classified according to its independent main verb behaviour. Note that the ‘matching’ is not merely a transitivity matching, but a matching at the level of category features under the kind of subevental decomposition advocated in this system.

In principle, one might expect constructivism to be a potentially powerful explanatory tool in dealing with a wide variety of complex predications, since it does not locate argument structure generalizations at the lexical level, but attempts to build them up compositionally from phrasal ingredients. However, the ‘ideal’ constructivist system would be monotonic, productive, structure building. Once intricate selectional facts come into play and insertion frames have to be stipulated, it is unclear whether one is really escaping from the reality of a lexical module with subcategorization frames and argument structure specification after all.

In the table below, I line up the predictions made from a pure ‘ideal’ structure building approach with the matching of category features approach. I assume for structure building that it should be possible to embed a verb root with a smaller set of arguments underneath an equal or larger ‘little v’ head, but not with a ‘smaller’ one. Thus, a transitive light verb should be able to sit on top of a root with either one or two arguments, and an intransitive light verb should only be able to combine with a root that has one argument.

(29)	MV	LV	Building	Match CAT features
		UNACC	✓	✓
	UNACC	UNERG	✓	X
		TRANS	✓	X
		UNACC	X	X
	UNERG	UNERG	✓	✓
		TRANS	✓	✓
		UNACC	X	X
	TRANS	UNERG	X	✓
		TRANS	✓	✓

As one can see, the category matching approach is the one that makes the right predictions in this case, pointing to the necessity of ‘selectional frames’ over and above the constructivist architecture.

In the next subsection, I lay out an implementation of the category matching facts in the modified constructivist account of Ramchand (2008b), exploiting the idea that roots are not category neutral but come specified with a set of category features that determine their behaviour. The crucial additional element of the analysis is that underassociated category features have to be licensed by AGREE.

### 3.2.1. An Analysis

Recall that in the discussion of lexicalization of articulated functional sequences, we assumed that there was a superset principle at work, and that while all functional structure had to be exhaustively lexicalized, category features on a particular lexical item could remain underassociated. Recall also that the superset principle thus stated was potentially too unconstrained, predicting general optionality of category features on any lexical item. To constrain the effects of Superset, I propose the following.

- (30) *Constraints on Underassociation:*
- (a) Underassociation of category features of any ‘main verb’ is in principle possible, constrained by AGREE.
  - (b) AGREE-ING categorial features must unify their conceptual content.

The idea here is that a verb is ‘light’ precisely when its categorial features are in an AGREE relation with another verbal element in its complement. Moreover, since the conceptual content of these AGREEING features must be unified, this will only be possible when one of the verbal elements has little, or extremely abstract and impoverished lexical content. The AGREE constraint will essentially force the underassociated features of the main verb to ‘match’ those of the light verb.

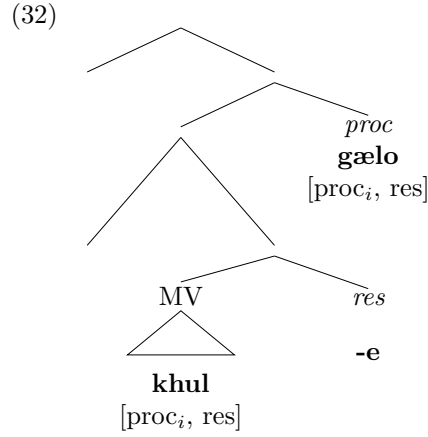
I make one further assumption for the particular case of the Bengali light verb constructions examined here, namely that the perfective ending carried by the main verb contributes a [*res*] feature to the construction.

With this in hand, we can show how the system operates in the cases of the different types of light verb examined in this section.

### Unaccusative LV

In the case of the unaccusative light verb, the light verb lexicalizes [*proc*], and the main verb lexicalizes [*res*] via the perfective ending with all its other features remaining underassociated.

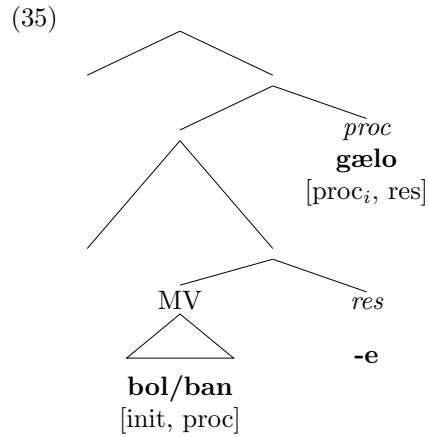
- (31) dorja-ṭa      khul-e    gælo *unaccusative*  
 door-CLASS. open-E go.PAST3  
 ‘The door opened.’



If the main verb contains an *[init]* feature, the structure will never converge because it will be licensed neither by direct lexicalization nor by AGREE.

- (33) \*O kothaṭa      bole    gælo *unergative*  
 He word-CLASS. speak-E go.PAST3  
 ‘He spoke’ (intended)

- (34) \*jon baṛiṭa      ban-iye    gælo *transitive*  
 john house-CLASS build-E go.PAST3  
 ‘John built the house.’ (intended)



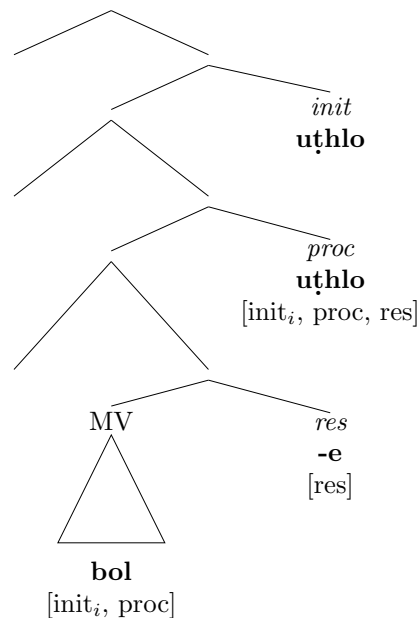


## Unergative LV

In the case of the unergative light verb, the light verb will lexicalize both *[init]* and *[proc]*, but now the underassociated main verb is free to have either *[proc]* or both *[init]* and *[proc]*, allowing all verb types in embedded position in principle.<sup>4</sup>

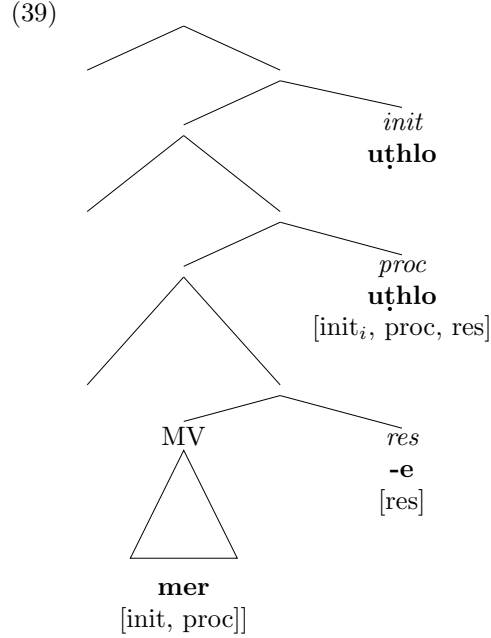
- (36) O bol-e uṭhlo *unergative*  
 he/she speak-E rise.PAST3  
 'He/she spoke up.'

- (37) (a)



- (38) O        bagh-ta-ke        lathi mer-e uṭhlo        *transitive*  
 he/she tiger-CLASS-ACC kick hit-E rise.PAST3  
 'He unexpectedly kicked the tiger.'

<sup>4</sup>I have no explanation for why transitive verbs are so rare under the unergative light verb, but I will tentatively assume that competition with the extremely productive transitive light verb *phæla*-‘drop’ is what is at stake here.



### Transitive LV

In the case of the transitive light verb, we are in a similar situation to the unergative light verb above. We predict therefore that all verbs will be possible embedded under transitives. The ungrammaticality of unaccusatives under the transitive light verb does not get a straightforward explanation in terms of AGREE of underassociated features, however. Here, we need to enforce matching in the other direction as well, and I propose that this is because of the constraint on using a verb as a ‘light verb’, which requires all of *its* features to be linked to another by means of AGREE.

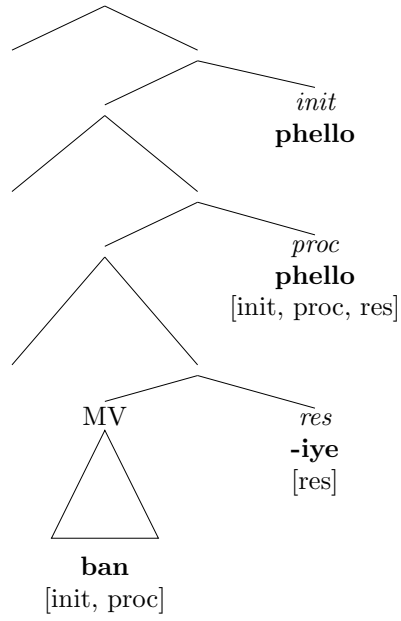
(40) *The Light Verb Constraint:*

A verb can be used as a light verb when all of its category features AGREE with some other verbal element in its complement domain

With this in hand, we get the restriction to just unergatives and transitives in the embedded position of the transitive light verbs.

- (41) O      baṛi-ṭa      ban-īye phello      *transitive*  
 he/she house-CLASS. build-E throw/drop.PAST3  
 ‘He accomplished/completed the building of the house.’

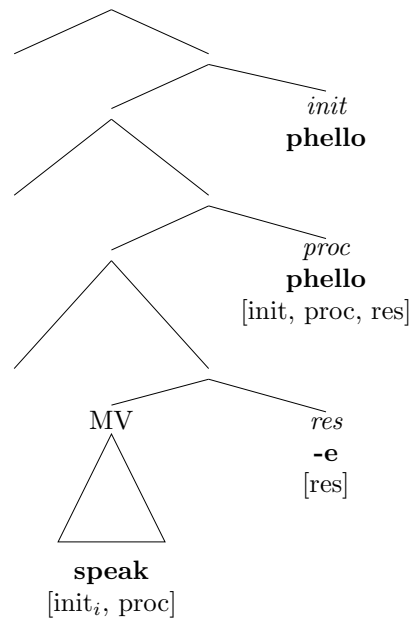
(42) (a)



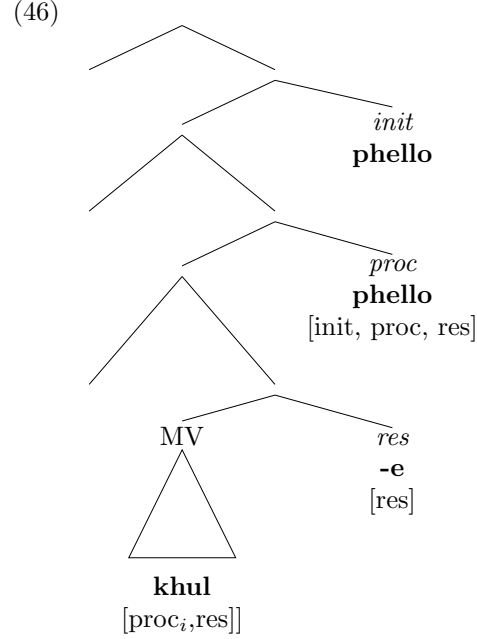
(43) ram bol-e phello  
ram speak-E throw/drop.PAST3  
'Ram blurted something out.'

*unergative*

(44)



- (45) \**dorja-ta khul-e phello* *unaccusative*  
 door-CLASS. open-E throw/drop.PAST3  
 ‘The door opened.’



Note that the light verb *Otha* ‘rise’ is able to combine with unaccusatives only because it also has an incarnation as an unaccusative verb. In its unergative instantiation, it too may not embed an unaccusative main verb, just like the transitive counterpart.

Some verbs in Bengali are used as light verbs and not others. For the purposes of this paper, I have stipulated that a light verb is one whose category features are all linked by AGREE to some other element in their complement. A natural question that arises at this point is whether this possibility has to be stipulated on individual verbs in the lexicon. Notating this possibility would be a serious addition to the information load carried by the lexicon, whereas I have been assuming that lexical information is confined to featural properties that are independently attested in the syntactic computation. One promising line of thought is to say that because of the constraint on unifying lexical conceptual content, AGREE of subevental features can only be achieved if the lexical conceptual content of the light verb is very abstract to begin with. This gets us quite far, but perhaps not far enough, since it is undeniable that these light verbs do have *more* lexical encyclopedic content when used on their own. On the other hand, the content they appear to have when used on their own is always in the physical/spatial domain. My speculation is that this content is never part of the lexical encyclopedia of the light verb at all, but filled in as a default

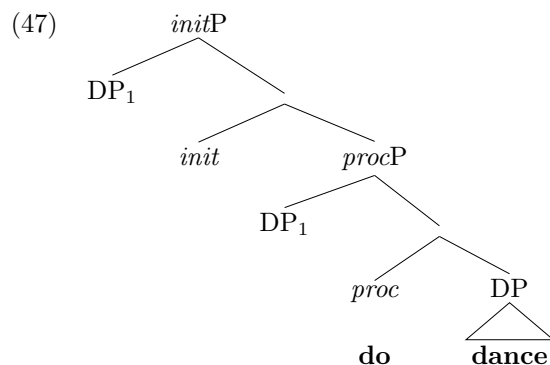
by the post-syntactic conceptual system. If this idea can be made to work, then light verbs will be precisely those verbs that have highly underspecified and abstract lexical encyclopedic content in the first place, and they do not need to be notated as allowing a ‘light verb’ use.

In the next section, I turn to another implementation of the idea of underassociation constrained by AGREE to reevaluate another class of verbal construction, this time in English, where selection seems to be in evidence.

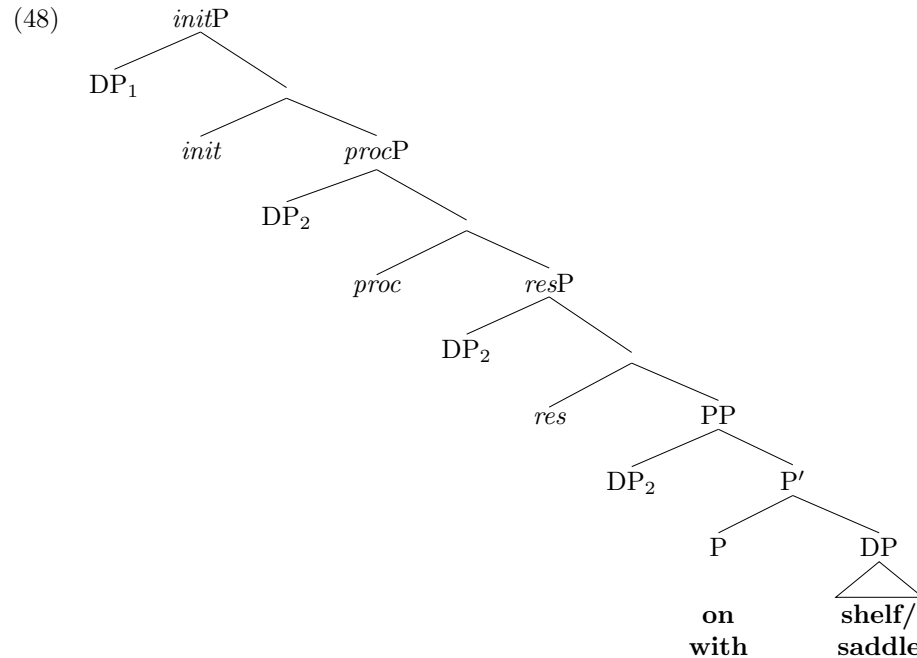
#### 4. The Problem of Denominal Verbs

In the lexical decomposition system of Hale and Keyser (1993) and subsequent work, a good deal of emphasis is placed on ‘conflation’ type verbs (denominal and deadjectival verbs) where it is claimed that the verb is derived by abstract incorporation into the head of the verbal projection from complement position, subject to principles of syntactic movement.

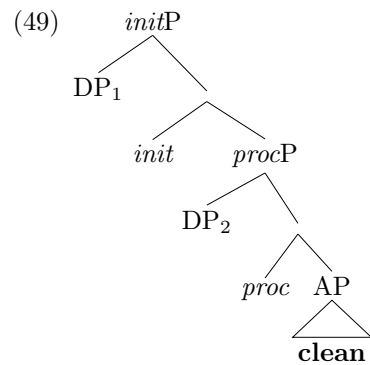
Conflation verbs seem to arise from rhematic material being incorporated from complement position into the head. In Hale and Keyser (1993), the verb *dance* is covertly transitive: the nominal ‘dance’ can be thought of as the complement of the generalized *do* process, which then conflates into the verbal head.



In the case of the location verbs, the nominal in question is the complement of the PP (what I would call RHEME of result, further describing the result state achieved by the undergoer of translational motion). In the case of locatum verbs, the RHEME of result is the possessional PP ‘with saddle’. So once again, the nominal ‘saddle’ is within the rhematic material of the clause, and incorporating it would be an (unproblematic) case of incorporation from a complement position.



In the case of deadjectival verbs, once again the incorporation seems to be from the AP rhematic complement of the *Res* head.



The intuition behind the Hale and Keyser account is that the correlation with selection (which determines the complement) and ‘conflation’ reflects a real syntactic generalization. In Hale and Keyser (2000), however, a distinction is made between conflation and genuine syntactic incorporation (which is assumed to be constrained by ‘government’). The problem is that ‘conflation’ verbs are compatible with an overt DP in complement position.

- (50) (a) They are dancing a Sligo jig.  
 (b) They shelved the books on the windowsill.  
 (Hale and Keyser 2000:49)

This compatibility forces Hale and Keyser to rephrase the issue in terms of ‘conflation’ which is now interpreted as a phonological copying distinct from genuine syntactic incorporation.

(51) *Conflation*:

Conflation consists in the process of copying the p-signature of the complement into the p-signature of the head, where the latter is ‘defective’.

(Hale and Keyser 2000:63)

The DP in object position is not free however, but seems to be confined to cognate objects or hyponymous objects (objects which denote a more specific or ‘precisified’ concept with respect to a more general superordinate nominal concept encoded in the verb). The following examples from Hale and Keyser illustrate the two types: cognate objects (52a, b) and hyponymous objects (52c, d).

(52) (a) She slept the sleep of the just.

(b) He laughed his last laugh.

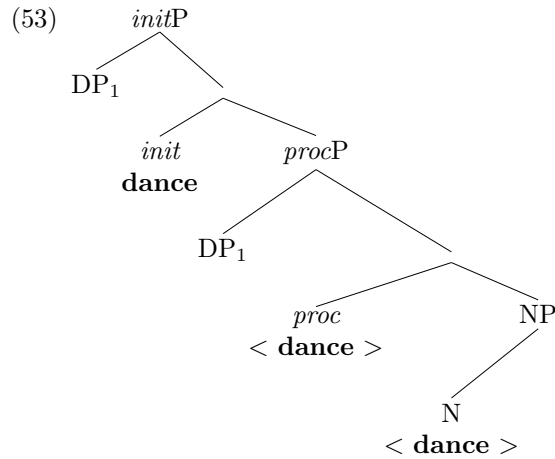
(c) He danced a jig.

(d) He bagged the potatoes in a gunnysack.

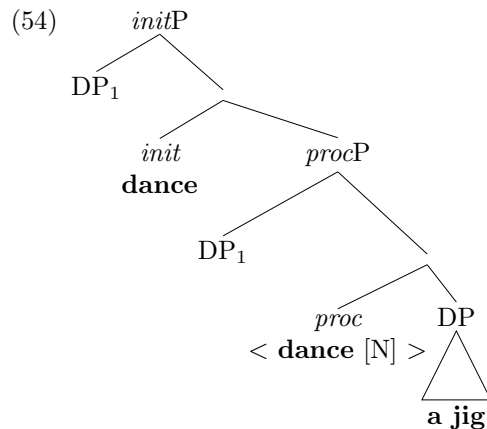
(Hale and Keyser 2000:71)

Hale and Keyser (2000) argue that the cognate objects are a true case of conflation, but where two copies of the p-signature are produced. Conflation is possible because they assume that the selectional relationship is not disrupted by functional elements of the extended projection of the head; both p-signatures are necessary because English prohibits the stranding of determiners. They leave it open whether the verbs with hyponymous objects should be given the same treatment or not.

However, the system of lexical insertion advocated in this paper offers us another way out of the ‘conflation’ verbs problem (assuming that we find reference to the notion of p-signature’ in a syntactic computation problematic). Since lexical items come with a clutch of category labels, and since the system abandons the assumption that lexical items are inserted under a single terminal node, it is possible to endow a lexical item like *sleep* or *dance* with a nominal feature in addition to its verbal features. If a root like *dance* is endowed with [init, proc, N] category features, it will be able to identify the subtree indicated, providing lexical encyclopedic content for not only the process but also the rhematic material of the process event.



The idea is that the structure in (53) above is not the only way to build a tree using the lexical item *dance*. In principle, the nominal feature of *dance* can underassociate, and an independent DP structure can be merged in the complement position (where the underassociated N feature on *dance* is shown in brackets).

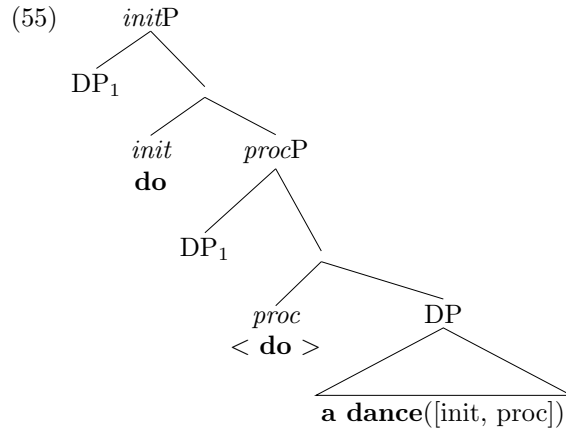


In the case at hand, denominal verbs, we must fill the complement with a projection that will identify N, and in addition, the information about the rheme provided by the conceptual content of the root (i.e. that the person is performing a ‘dance’) must be unified with the conceptual content of the DP complement. This will be successful if ‘dance’ and the denotation of the DP stand in a hyponymous relation (and by this I include the case of the so called ‘cognate’ objects, since I no longer need to rely on identity of p-signatures).

English is special in this regard, reduced to a fact about large chunks of its lexical inventory which contains items with both nominal and verbal



features. This fact, together with the superset principle, means that a given lexical item in English can be used in both nominal and verbal environments provided its underassociated features can be otherwise satisfied. Thus, in a complex predicate, if the light verb *do* inserts to identify process, then *dance* may underassociate to identify just the nominal complement part of the structure. Here, the underassociated [init, proc] features of *dance* will have to unify with the information about process and initiation provided by *do*. This immediately accounts for why the structure is only possible when the verbal identifier (here *do*) is fairly underspecified for conceptual content, what we generally call a ‘light’ verb.



Note that clause two of the Underassociation principle virtually ensures that lexical content of helping verbs in complex predicate constructions once again, will be abstract and generally underspecified as we found with the Bengali construction in the previous section. Hence the term ‘light’ verbs. These are the only verbs that will be hyponymous with their host contentful verbal roots.<sup>5</sup>

In short, I would argue that a more satisfying implementation of conflation is thus achieved within this set of assumptions, without extra architectural mechanisms.

## 5. Conclusion

Going back to the notion of selection, we see that in addition to (2ai) (the ordering imposed by extended projection/functional sequence, we can actually eliminate (2aii) altogether, together with its special rider concerning first merge.<sup>6</sup> Instead, we make good on assumption three, which stated

<sup>5</sup>I have nothing to say here about the pure DP uses of the lexical item *dance* in English. I assume that some head in the nominal functional sequence is responsible for binding off or otherwise neutralizing the verbal features of *dance* in these cases.

<sup>6</sup>Note that the stipulation about first merge is unnecessary because of the requirement that a vocabulary item spell out a contiguous sequence of heads.

that ‘selection can be achieved via category labels on lexical items combined with a constrained theory of lexical insertion. Doing so involved actually proposing a constrained theory of lexical insertion, which incorporated the following principles.

(56) *The Superset Principle*

The phonological exponent of a Vocabulary item is inserted into a node if the item matches all or a superset of the grammatical features specified in the node. Insertion does not take place if the Vocabulary item does not contain all features present in the node. Where several Vocabulary items meet the conditions for insertion, the item containing fewer features unspecified in the node must be chosen.

(57) *Exhaustive Lexicalization*

Every node in the syntactic representation must be identified by lexical content.

(58) *Underassociation:*

If a lexical item contains an underassociated category feature,  
 (i) that feature must be independently identified within the phase and linked to the underassociated feature, by Agree;  
 (ii) the two category features so linked must unify their lexical encyclopedic content.

Underassociation, as constrained by AGREE is doing the job of selectional features in the two cases that I have examined in this paper. If this analysis is on the right track, then the mechanisms discussed here could potentially be invoked to account for a number of other cases of c-selection described in the literature, thus eliminating the distinction between inherent features and selectional features in the syntactic computation.<sup>7</sup> I leave investigation of this issue to further research.

## References

- Adger, David. 2003. *Core Syntax*. Oxford University Press, Oxford.  
 Borer, Hagit. 2005. *Structuring Sense: An Exo-Skeletal Trilogy*. Oxford University Press, New York.  
 Butt, Miriam. 1995. *The structure of complex predicates in Urdu*. Dissertations in linguistics. CSLI Publications, Stanford, Calif.

---

<sup>7</sup>Peter Svenonius (p.c.) points out that the option of underassociating all the ‘higher’ verbal features of *dance* in the light verb construction ‘do a dance’ potentially overgenerates, if all c-selection is to be handled by the underassociation under AGREE of a single ‘lower’ category feature. It is probable that the light verb option that *dance* participates in is an extra and not generally available strategy whose stricter conditions of possibility I have not yet understood.

- Butt, Miriam. 2003. The morpheme that wouldn't go away. Handout, University of Manchester seminar series.
- Butt, Miriam and Aditi Lahiri. 2005. Historical stability vs. historical change. Ms., University of Konstanz.
- Caha, Pavel. 2007. The superset principle. Unpublished ms.,.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford University Press, New York.
- Emonds, Joseph E. 2000. *Lexicon and Grammar: The English Syntacticon*. Mouton de Gruyter, Berlin and New York.
- Fábregas, Antonio. 2007. The exhaustive lexicalization principle. *Nordlyd Tromsø Working Papers in Linguistics* 34 2.
- Hale, Ken and Jay Keyser. 2000. On the time of Merge.
- Hale, Kenneth and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by Kenneth Hale and Samuel Jay Keyser, no. 24 in Current Studies in Linguistics, pp. 53–109. MIT Press, Cambridge, Ma.
- Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by Kenneth Hale and Samuel Jay Keyse, pp. 111–176. MIT Press, Cambridge, Ma.
- Harley, Heidi and Rolf Noyer. 1999. State-of-the-article: Distributed morphology. *GLoT* 4.4: 3–9.
- Marantz, Alec. 2001. Words. Ms. MIT.
- Paul, Soma. 2003. Composition of compound verbs in bangla. *Proceedings of the Workshop on MultiVerb Constructions, Trondheim Summer School*.
- Ramchand, Gillian. 2008a. Complex predicates in Bengali. Ms., data and classification.
- Ramchand, Gillian. 2008b. *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge University Press, Cambridge, UK.
- Ramchand, Gillian and Peter Svenonius. 2002. The lexical syntax and lexical semantics of the verb-particle construction. In *Proceedings of WCCFL 21*, edited by Line Mikkelsen and Chris Potts. Cascadilla Press, Somerville, Ma.
- Talmy, Leonard. 1978. Figure and ground in complex sentences. In *Universals of Human Language*, edited by Joseph H. Greenberg, vol. 4, pp. 625–649. Stanford University Press, Stanford, Ca.